Chapter 8

Public Attitudes Toward Data Sharing by Federal Agencies

Eleanor Singer and John VanHoewyk, University of Michigan
Stanley Presser, University of Maryland

Abstract

Very little information exists concerning public attitudes on the topic of data sharing among Federal agencies. The most extensive information prior to 1995 comes from questions on several IRS surveys of taxpayers, from questions added to a series of Wisconsin surveys carried out in 1993-95, and from scattered other surveys reviewed by Blair (1995) for the National Academy of Sciences panels. From this review it is clear that the public is not well informed about what data sharing actually entails, nor about the meaning of confidentiality. It seems likely that opinions on this topic are not firmly held and liable to change depending on other information stipulated in the survey questions as well as on other features of the current social climate.

In the spring of 1995, the Survey Research Center at the University of Maryland (JPSM) carried out a random digit dialing (RDD) national survey which was focused on the issue of data sharing. The Maryland survey asked questions designed to probe the public's understanding of the Census Bureau's pledge of confidentiality and their confidence in that pledge. Respondents were also asked how they felt about the Census Bureau's obtaining some information from other government agencies in order to improve the decennial count, reduce burden, and reduce cost. In addition, in an effort to understand responses to the data sharing questions, the survey asked about attitudes toward government and about privacy in general.

Then, in the fall of 1996, Westat, Inc. repeated the JPSM survey and, in addition, added a number of split-ballot experiments to permit better understanding of some of the responses to the earlier survey. This paper examines public attitudes toward the Census Bureau's use of other agencies' administrative records. It analyzes the relationship of demographic characteristics to these attitudes as well as the interrelationship of trust in government, attitudes toward data sharing, and general concerns about privacy. It also reports on trends in attitudes between 1995 and 1996 and on the results of the question-wording experiments imbedded in the 1996 survey. Implications are drawn for potential reactions to increased use of administrative records by the Census Bureau.

Introduction

For a variety of reasons, government agencies are attempting to satisfy some of their needs for information about individuals by linking administrative records which they and other agencies already possess. Some of the reasons for record linkage have to do with more efficient and more economical data collection, others with a desire to reduce the burden on respondents, and still others with a need to improve coverage of the population and the quality of the information obtained.

The technical problems involved in such record linkage are formidable, but they can be defined relatively precisely. More elusive are problems arising both from concerns individuals may have about the con-

fidentiality of their information and from their desire to control the use made of information about them. Thus, public acceptance of data sharing among Federal and state statistical agencies is presumably necessary for effective implementation of such a procedure, but only limited information exists concerning public attitudes on this topic.

A year and a half ago, the Joint Program in Survey Methodology (JPSM) at the University of Maryland devoted its practicum survey to examining these issues. The survey asked questions designed to probe the public's understanding of the Census Bureau's pledge of confidentiality and their confidence in that pledge. It also asked how respondents felt about the Census Bureau's obtaining some information from other government agencies in order to improve the decennial count or to reduce its cost. In addition, in an effort to understand responses to the data sharing questions, the survey asked a series of questions about attitudes toward government and about privacy in general.

Most of these questions were replicated in a survey carried out by Westat, Inc. in the fall of 1996, a little more than a year after the original survey. The Westat survey asked several other questions in addition -- questions designed to answer some puzzles in the original survey, and also to see whether the public was willing to put its money where its mouth was -- i.e., to provide social security numbers (SSN's) in order to facilitate data sharing. Today, I will do four things:

- Report on trends in the most significant attitudes probed by both surveys;
- Discuss answers to the question about providing social security numbers;
- Report on progress in solving the puzzles left by the JPSM survey; and
- Discuss the implications of the foregoing for public acceptance of data sharing by Federal agencies.

Description of the Two Surveys

The 1995 JPSM survey was administered between late February and early July to a two-stage Mitofsky-Waksberg random digit dial sample of households in the continental United States. In each household, one respondent over 18 years of age was selected at random using a Kish (1967) procedure. The response rate (interviews divided by the total sample less businesses, nonworking numbers, and numbers that were never answered after a minimum of twenty calls) was 65.0 percent. The nonresponse consisted of 23.4% refusals, 6.5% not-at-home, and 5.1% other (e.g., language other than English and illness). Computer-assisted telephone interviewing was conducted largely by University of Maryland Research Center interviewers, supplemented by graduate students in the JPSM practicum (who had participated in the design of the questionnaire through focus groups, cognitive interviews, and conventional pretests). The total number of completed interviews was 1,443.

The Westat survey (Kerwin and Edwards, 1996) was also conducted with a sample of individuals 18 or older in U.S. households from June 11 to mid-September. The response rate, estimated in the same way as the JPSM sample, was 60.4% [1]. The sample was selected using a list-assisted random digit dial method. One respondent 18 or over was selected at random to be interviewed.

Trends in Public Attitudes Toward Data Sharing

The most significant finding emerging from a comparison of the two surveys was the absence of change with respect to attitudes relating to data sharing. Indeed, if we are right that there has been little change on these matters, the new survey is testimony to the ability to measure attitudes reliably when question wording, context, and procedures are held reasonably constant -- even on issues on which the public is not well informed and on which attitudes have not crystallized. In 1996 between 69.3% and 76.1%, depending on the agency, approved of other agencies sharing information from administrative records with the Census Bureau in order to improve the accuracy of the count, compared with 70.2% to 76.1% in 1995 [2]. Responses to the Immigration and Naturalization Service, asked about in 1995, and the Food Stamp Office, asked about in 1996, are comparable to those to the Social Security Administration (SSA). Responses are consistently least favorable toward the Internal Revenue Service (IRS).

Westat documents five significant changes ($p < .10$) among 22 questions asked about the Census Bureau on both surveys. First, there is more awareness of the fact that census data are used to apportion Congress and as a basis for providing aid to communities; but second, there is less awareness that some people are sent the long census form instead of the short form. (Both of these changes make sense in retrospect. In the election year of 1996, apportionment was very much in the news; at the same time, an additional year had elapsed since census forms, long or short, had been sent to anyone.) Third, fewer people in 1996 than 1995 said that the five questions asked on the census short form are an invasion of privacy -- a finding at odds with others, reported below, which suggest increasing sensitivity to privacy issues between the two years. This issue will be examined again in the 1997 survey. Fourth, there was a modest increase in the strength with which people opposed data sharing by the IRS. This finding (not replicated with the item about data sharing by SSA) may have less to do with data sharing than with increased hostility toward the IRS. These changes are mostly on the order of a few percentage points. Finally, among the minority who thought other agencies could not get identifiable Census data there was a substantial decline in certainty, although the numbers of respondents being compared are very small.

Trends in Attitudes Toward Privacy

In contrast with attitudes toward data sharing and the Census Bureau, which showed virtually no change between 1995 and 1996, most questions about privacy and alienation from government showed significant change, all in the direction of more concern about privacy and more alienation from government. The relevant data are shown in Table 1.

There was a significant decrease in the percentage agreeing that "people's rights to privacy are well protected" and a insignificant increase in the percentage agreeing that "people have lost all control over how personal information about them is used." At the same time, there was a significant decline in the percentage *dis*agreeing with the statement, "People like me don't have any say about what the government does," and a significant increase in the percentage agreeing that "I don't think public officials care much what people like me think" and in the percentage responding "almost never" to the question, "How much do you trust the government in Washington to do what is right?" The significant decline in trust and attachment to government manifested by these questions is especially impressive given the absence of change in responses to the data sharing questions. We return to the implications of these findings in the concluding section of the paper.

Table 1. -- Concerns about Privacy and Alienation from Government, by Year

Attitude/Opinion	Agree Strongly or Somewhat	

	1995		1996	
People's rights to privacy are well protected	41.4	(1,413)	37.0	(1,198)
People have lost all control over how personal information about them is used	79.5	(1,398)	80.4	(1,193)
People like me don't have any say about what the government does	59.2	(1,413)	62.9	(1,200)
I don't think public officials care much what people like me think	65.4	(1,414)	71.1	(1,202)
How much do you trust the government in Washington to do what is right? (Almost never)	19.2	(1,430)	25.0	(1,204)

Willingness to Provide Social Security Number to Facilitate Data Sharing

One question of particular importance to the Census Bureau is the extent to which people would be willing to provide their social security number to the Census Bureau in order to permit more precise matching of administrative and census records. Evidence from earlier Census Bureau research is conflicting in this regard. On the one hand, respondents in four out of five focus groups were overwhelmingly opposed to this practice when they were asked about it in 1992 (Singer and Miller, 1992). On the other hand, respondents to a field experiment in 1992 were only 3.4 percentage points less likely to return a census form when it requested their SSN than when it did not; an additional 13.9 percent returned the form but did not provide a SSN (Singer, Bates, and Miller, 1992).

To clarify this issue further, the Bureau asked Westat to include a question about SSN on the 1996 survey. The question (Q21) read as follows:

> *"The Census Bureau is considering ways to combine information from Federal, state, and local agencies to reduce the costs of trying to count every person in this country. Access to social security numbers makes it easier to do this. If the census form asked for your social security number, would you be willing to provide it?"*

About two thirds (65.9%) of the sample said they would be willing to provide the number; 30.5% said they would not; and 3.5% said don't know or did not answer the question.

The question about SSN was asked *after* the series of questions asking whether or not people approved of other administrative agencies sharing data with the Census Bureau. Therefore, it is reasonable to assume that responses to this question were influenced by opinions about data sharing, which the preceding questions had either brought to mind or helped to create. And, not surprisingly, there is a relationship between a large number -- but not all -- of the preceding questions and the question about providing one's SSN.

For example, those who would provide their SSN to the Bureau are more likely to believe the census is extremely or very important and more likely to be aware of census uses. They are more likely to favor data sharing. Those who would not provide their SSN to the Bureau are more concerned about privacy issues.

They are less likely to trust the Bureau to keep census responses confidential; they are more likely to say they would be bothered "a lot" if another agency got their census responses; they are less likely to agree that their rights to privacy are well protected; less likely to believe that the benefits of data sharing outweigh the loss of privacy this would entail, and more likely to believe that asking the five demographic items is an invasion of privacy. All of these differences are statistically significant.

Table 2. -- Willingness to Provide SSN and Attitudes to Census Bureau

Attitude/Opinion	Would Not Provide SSN %	Would Provide SSN %
Believes counting population is "extremely" or "very" important	63.8	79.7
Is aware of census uses	43.1	54.8
Would favor SSA giving Census Bureau short-form information	56.3	85.0
Would favor IRS giving Census Bureau long-form information	30.4	61.2
Would favor "records-only" census	45.6	60.0
Trusts Bureau to not give out/keep confidential census responses	45.0	76.7
Would be bothered "a lot" if other agency got census responses	54.1	29.9
Believes benefits of record sharing outweigh privacy loss	36.0	51.1
Believes the five items on short form are invasion of privacy	31.3	13.4

There are also significant relationships between political efficacy, feelings that rights to privacy are well protected, feelings that people have lost control over personal information, and trust in "the government in Washington to do what is right" (Q24a-d) and willingness to provide one's SSN. These political attitude questions, it should be noted, were asked ***after*** the question about providing one's SSN, and so they could not have influenced the response to this question.

Of the demographic characteristics, only two -- gender and education -- are significantly (for gender, $p<.10$; for education, $p<.05$) related to willingness to provide one's SSN. Almost three quarters (71.4%) of men, but only 65.5% of women, are willing to provide their SSN. This is true of 71.2% of those with less than a high school education, 63.9% of those who are high school graduates, 68.7% of those with some college, and 76.8% of those who are college graduates. The same curvilinear relationship is apparent for income: 75.4% of those with family incomes of less than $20,000, 69.6% of those with incomes between $20,001 and $30,000 and $30,001 and $50,000, 68.6% of those between $50,001 and $75,000, and 75.4% of those with incomes over $75,000 say they would be willing to provide their SSN if asked by the Census Bureau to do so.

Table 3. -- Willingness to Provide SSN, by Concerns about Privacy and Alienation from Government

Concern/Alienation	Would Provide SSN %	Would Not Provide SSN %
Disagrees strongly that rights to privacy are well protected	24.2	45.6
Agrees strongly people have lost control over personal information	37.9	54.2
Agrees strongly "people like me" have no say about what government does	27.7	43.7
Agrees strongly public officials don't care much about "what people like me think"	31.2	45.4
Almost never trusts government in Washington to do what's right	19.5	37.8
Privacy loss outweighs economic benefit of data sharing	47.1	56.0
Economic benefit of data sharing outweighs privacy loss	47.9	30.4

From the foregoing, it appears that there are two reasons underlying reluctance to provide one's SSN. *First, there are reasons associated with beliefs about the census:* People who are less aware of the census, who consider it less important, and who are less favorable toward the idea of data sharing are significantly less willing to provide their SSN. Low levels of education are also associated with these characteristics. *Second, however, is a set of beliefs and attitudes concerning privacy, confidentiality, and trust:* People who are more concerned about privacy, who have less trust in the Bureau's maintenance of confidentiality, and who are less trusting of government in general are much less likely to say they would provide their SSN to the Census Bureau. Women are more likely to be concerned about privacy issues than men, and they are also less willing to say they would provide their SSN to the Bureau. In earlier analyses (Singer and Presser, 1996) we found that importance attached to the census, knowledge about the census, and attitudes about privacy were independent factors predicting willingness to have other agencies share data with the Bureau. Though we have not carried out a factor analysis of attitudes toward willingness to provide one's SSN, the relationships described above suggest that the same clusters of beliefs are relevant for this attitude, as well.

We should point out that the question asked on the 1996 survey, about whether or not respondents would be willing to provide their SSN, is not equivalent to a field experiment. The number of people who would provide their SSN if asked to do so in an actual census might very well be higher than the two thirds who said they would do so on this survey, as suggested by the field experiment cited at the beginning of this section. On the other hand, if the issue of privacy became salient prior to the census, the number complying might well be less. Arguing for the second, more cautious, inference is the fact that more than a third of those approached for the survey did not participate, and, since the introduction to the survey informed potential respondents about the topic, the nonparticipants may well have included those more suspicious of government and less inclined to cooperate with any request from government agencies, including the Census Bureau [3].

What Does Confidentiality Mean?

A number of question wording experiments were included in the 1996 Westat survey. The most important of these, from the perspective of understanding data sharing attitudes, had to do with the meaning of the Census Bureau's assurance of data confidentiality to respondents. The short answer to the question, “What does confidentiality mean to the public?” is, “We don't know.” However, in the rest of this paper, we try to summarize what we think we learned.

The 1995 JPSM survey resulted in one very puzzling finding. When asked whether other agencies could get their answers to census questions, *identified by name and address*, 41% said they did not know; of the rest, about 90% said other agencies could get such information (Presser and Singer, 1995). To make things even more puzzling, the better educated were more likely to believe, erroneously, that other agencies could get such data -- virtually the only time, so far as we know, that more education has been associated with more error (Hyman, Wright, and Reed, 1975). Furthermore, the belief that other agencies *could* get such data was associated with *more favorable* attitudes toward data sharing.

It thus seemed fairly clear that our attempt to provide a neutral definition of "confidentiality" in the 1995 instrument had not had the intended effect. Accordingly, we incorporated a four-way split ballot experiment into the 1996 survey.

One quarter of the sample were asked the 1995 question; one quarter, the 1995 question without the DK filter. One quarter were asked, "Do you think the Census Bureau does or does not protect the confidentiality of this (household demographic) information, or don't you know (DK)?" And the final quarter were asked the confidentiality question without the DK filter.

The results are shown in Table 4. The most striking thing about the table is simply the variation in responses, depending on the wording of the question. But the next most startling finding is the difference in responses to the questions asking whether other agencies can get identified data, and whether the Bureau keeps data confidential. Omitting those who answer DK, the percentages who believe responses are NOT shared (or data ARE kept confidential) ranges from 11.5% in Q 7-1 to 69.2% in Q 7-4. Omission of the DK filter reduces the size but does not change the basic form of the relationship. Majorities of the public believe that other agencies can get identified data; they also believe that the Bureau maintains data confidentiality.

Table 4. -- The Effects of Question Wording on Beliefs Regarding Sharing of Responses by Census Bureau

Response	Do you think other government agencies...can or cannot get people's names and addresses along with their answers to the census?		Do you think the Census Bureau does or does not protect the confidentiality of this [household demographic] information?	
	Explicit "Not Sure" %	No Explicit "Not Sure" %	Explicit "Not Sure" %	No Explicit "Not Sure" %
Believe that census responses are shared	47.1	76.9	9.6	20.9
Believe that census responses are not shared	6.1	15.4	12.9	47.0
Not Sure/ Don't Know	46.8	7.7	77.5	32.1
N (unweighted)	310	296	294	315

In passing, we should note that the distribution of answers to the version of the question which is identical to the 1995 question do not differ significantly from the 1995 distribution; and, as in 1995, people who said other agencies CAN get data were significantly more likely to favor data sharing in 1996 as well.

In another effort to understand the meaning of confidentiality to respondents, we asked another split-ballot question near the end of the 1996 survey. One asked whether the Census Bureau was required by law to keep census information confidential; the other, whether the Bureau was forbidden by law from giving identified census information to other agencies.

The responses to the two versions of this question are shown in Table 5. Majorities of those who have an opinion give the correct answer to both questions; but the proportion answering DK is larger, and the proportion giving the correct answer smaller, when the question asks about giving other agencies identified information than when it asks about maintaining confidentiality.

As a follow-up to both questions, we asked those who said the Bureau is required to protect the information or forbidden from disclosing it, *whether or not they trusted the Bureau to uphold the law -- that is, to keep the information confidential, or to refrain from disclosing it to other agencies.* Regardless of which version of Q22 they got, two thirds of those who answered Yes to the factual question about legal requirements said they trusted the Bureau to comply with the law. However, *those who not only say the Bureau is required to keep information confidential but who also trust the Bureau to do so, are significantly more likely to say both that other agencies cannot get the data and that the Bureau keeps data confidential.* Thus, not only knowledge of the law, but also trust in the Bureau's compliance with the law, is implicated in responses to the factual questions about whether the Bureau does or does not protect the data in its possession.

Table 5. -- The Effect of Question Wording on Knowledge of Laws Regarding Sharing of Census Information

Response	Is the Census Bureau forbidden by law from giving other government agencies census information identified by name or address?	Is the Census Bureau required by law to keep census information confidential?	Total
	%	%	%
Yes	28.3	51.1	40.2
No	17.1	11.6	14.2
Dont't Know	54.6	37.3	45.5
N (unweighted)	591	624	121

5

What differentiates those who trust the Bureau to keep information confidential from those who do not?

We found only two demographic characteristics that seemed to make a difference. Women are considerably more likely to say they trust the Bureau than men, and younger respondents are more likely to express trust than older respondents are. Whether this is an effect of age or of cohort is impossible to tell from this cross-sectional survey. None of the other demographic characteristics we examined -- education, race, or income -- make a consistent difference in attitudes of trust.

Finally, we looked at the relation of the beliefs about legal requirements to attitudes about data sharing. People who believe the Bureau is required by law to keep data confidential are significantly more likely to favor data sharing than those who do not. On the other hand, people who believe the Bureau is forbidden from sharing data *with* other agencies are significantly more likely to oppose data sharing *by* other agencies. Whether this results from confusion, or from an application of the norm of reciprocity, or from opposition to all data sharing, is impossible to tell.

Conclusions and Implications

The following conclusions seem to follow from comparison of the 1995 and 1996 surveys:

- Beliefs about the Census Bureau and attitudes toward data sharing have undergone little change since 1995.
- Beliefs about privacy and trust in government have deteriorated since 1995.
- To the public, the belief that the Bureau protects confidentiality does not seem to mean that other agencies cannot get data identified by name and address. What it does mean, we cannot tell from these data.
- In contrast to an implicit Census Bureau hypothesis, knowledge about legal requirements for confidentiality is not enough to convince the public that the Bureau actually protects confidentiality. In

order for knowledge to translate into belief, trust in the Bureau is required. The number of people who both know about legal requirements and trust the Bureau is only about two thirds as great as the number whose factual information is correct. (However, both knowledge and trust are independently related to attitudes toward data sharing.)

We believe these findings have two major implications for future data collection:

- First, we are planning in 1997 to ask both about whether other agencies can get data, and whether the Census Bureau maintains confidentiality, of one third of the sample. Then, everyone will be asked what confidentiality means to them. Only when the sources of misunderstanding are known can the Bureau better communicate its message about data protection to the public.

- Second, future surveys should be used to experiment with arguments that might be presented to the public in favor of data sharing. For example, there is evidence from the 1995 and 1996 surveys that the quality of the data is a more important consideration than cost. Are there orther arguments that are even more persuasive? How can the argument about quality be made even more compelling?

We hesitate to make substantive predictions about the public's acceptance of data sharing at the time of the next census. On the one hand, about two thirds of the public currently to favor this practice, this proportion has remained stable over at least a year, and two thirds say they would be willing to provide their SSN to the Bureau to facilitate such sharing. On the other hand, opposition to data sharing, and to making the SSN available, is strongly related to privacy concerns, and such concerns show a small but significant increase between 1995 and 1996. Thus, it seems possible that if privacy concerns continue to increase, they may erode the support for data sharing that currently exists. The same implication can be drawn from our findings concerning belief in the Census Bureau's assurance of confidentiality. Information about the law is apparently not enough; trust is also required. And the latter is a much more difficult message to communicate effectively.

Acknowledgments

We would like to express our thanks to Jeff Kerwin and Sherman Edwards at Westat for help with some aspects of data analysis and to Randall Neugebauer and his colleagues at the Census Bureau for helpful comments on an earlier draft.

Note

Follow-up to this research appears in Singer, Eleanor and Presser, Stanley (1997). Public Attitudes Toward Data Sharing by Federal Agencies: Trends and Implications, *Survey Research Methods Section Proceedings*, American Statistical Association (in process).

Footnotes

[1] The Westat report gives a response rate of 64.4%, which is based on excluding the number of respondents with language problems (n=126) from the denominator. This group is included in the JPSM count of eligibles. The introduction to the Westat survey differed somewhat from that used by JPSM. It read as follows:

"My name is _____. I'm calling from Westat on behalf of the U.S. Census Bureau in Washington, D.C. We're doing a study of people's opinions on whether government agen-

cies keep information about them private. You were randomly selected for this study from the adults in your household. This survey has been approved by the Office of Management and Budget, Number 0607-0822. Without this approval, we could not conduct this survey. Any questions or comments about the survey may be directed to the Census Bureau. If you would like, when we are done, I will provide you with the address."

The JPSM introduction omitted all references to OMB or the Census Bureau, as well as the sentence about random selection, and introduced the interviewer as calling from the University of Maryland. The sentence about the topic of the study was identical to that in the Westat introduction.

[2] Text and tables use data weighted for number of residential phone numbers in the household and number of persons in the household, poststratified to Census estimates of sex, race, age, education, and region.

[3] If the Bureau used a less specific introduction, the overall response rate to the survey might not change, but nonrespondents might be more representative (less biased) with respect to their attitudes toward government and the Census Bureau.

References

Blair, Johnny (1995). Ancillary Uses of Government Administrative Data on Individuals: Public Perceptions and Attitudes, a paper commissioned by the panel to Evaluate Alternative Census Methods, National Academy of Sciences, College Park, MD: Survey Research Center, University of Maryland.

Hyman, Herbert H.; Wright, Charles; and Reed, John (1975). *The Enduring Effects of Education,* Chicago: University of Chicago Press.

Kerwin, Jeffrey and Edwards, Sherman (1996). The 1996 Survey on Privacy and Administrative Records Use, Staff paper, Rockville, MD: Westat, Inc.

Kish, L. (1967). Survey Sampling, New York: Wiley.

Presser, Stanley and Singer, Eleanor (1995). Public Beliefs about Census Confidentiality, paper presented at the 1995 meetings of the American Sociological Association.

Singer, Eleanor and Miller, Esther R. (1992). Report on Focus Groups, Center for Survey Methods Research, U.S. Bureau of the Census.

Singer, Eleanor; Bates, Nancy; and Miller, Esther R. (1992). Memorandum for Susan Miskura, Bureau of the Census, July 15, 1992.

Singer, Eleanor and Presser, Stanley (1996). Public Attitudes toward Data Sharing by Federal Agencies, paper presented at the Annual Research Conference, Bureau of the Census.